

Contents

Foreword	vii
Preface	vii
Acknowledgments	x
Chapter 1: In Search of a Visually Based Language System	1
Introduction	1
Language and Conversation Skills	1
Learning Capacity and Comprehension	2
Ability to Refer to and Orient to Time	2
Developments in the Use of Visuals	5
References	11
Chapter 2: Fundamentals of Communication and Language Acquisition	13
Symbols	13
Language Fundamentals	19
References	24
Chapter 3: Three Modes of Visual Language: VEM, VOM, and VIM	25
Visual Cues	25
Communication Displays	27
The Three Modes	34
Does Visual Language Blunt Speech?	41
References	43
Chapter 4: Assessment, Part 1: Visual Skills and Other Key Skills	45
Obstacles to Assessment	46
Testing Strategies and Key Assessment Principles	49
Assessment for Placement in Visual Immersion Program	52
References	65
Chapter 5: Assessment Part 2: Experiential Knowledge	67
Introduction	67
Previous Research	69
The Experiential Knowledge Profile	70
References	77

Chapter 6: Intervention: Applying the Three Modes of the VIP	79
Introduction	79
General Instructional Objectives	79
Areas of Communication Skill Development	81
Additional Intervention Principles and Guidelines	83
Language Instruction Mapped to Routines	87
Visual Immersion Program Instruction	90
Temporal Displays	106
References	109
Chapter 7: Special Considerations	111
Controlling Inappropriate Behavior	111
Expression of Pain and Discomfort	115
The Inclusive Classroom	115
Using Visuals Inside and Outside the Home	117
Interventions for Pre-Level 1 Learners	118
Intervention for Learners Who Read	120
Selecting Appropriate Assistive Technology	123
References	126
Appendix A: Monarch Natural Language Assessment	127
Appendix B: Informal Language Milestones	131
Appendix C: Experiential Knowledge Profile (EKP)	133
Appendix D: Monarch Individualized Pain, Illness, and Discomfort Awareness Program	159
Appendix E: Key Factors in Selecting an Electronic Communication Device	161
Index	171

3

Presenting Visuals in the Three Modes of the VIP

The use of visual symbols to help learners with ASD improve their communication skills has been an important component of clinical work since the 1980s. However, symbols are often applied inconsistently and without an overall approach. A systematic application of visuals with a prescribed scope and sequence based on the learner's strengths and weaknesses is more likely to lead to increased comprehension, speech production, language growth, and overall improvement in behavior.

Our program offers a comprehensive model based on using visuals in three modes—expression, instruction, and organization. The three modes use a range of visual symbols in a variety of ways. Before examining the modes themselves, we first need to review the ways visual symbols can be presented and the types of communication displays in which they are used.

Visual Cues

Visual symbols can be presented in two ways, as element cues or scene cues.

Element Cues

An element cue is a symbol that illustrates an individual component within a scene. There are six types used within VIP:

- **Agent element cues**—symbols that represent people or characters (conventionally referred to as nouns)
- **Object element cues**—symbols that represent things that can be seen, touched, or heard (also nouns)
- **Action element cues**—symbols that represent activity (verbs)

- **Spatial element cues**—symbols that indicate the position of one object with respect to another (prepositions)
- **Attribute element cues**—symbols that describe a key quality or characteristic of an agent, action, or object (adjectives)
- **Temporal element cues**—symbols that indicate duration or frequency (generally are adjectives or adverbs).

Figure 3-1 provides examples of the different types of element cues.

Element cues can be presented singly or in combined form. For example, a picture of a dog (an agent element cue) is a single element cue. A picture of a dog presented simultaneously with a picture of a brown circle (attribute element cue) and a picture of legs running (action element cue) is a combined element cue that represents a brown dog running. Combined element cues provide detailed information about a situation or event, but require greater understanding of semantic relationships from the learner.

Scene Cues

A scene cue is a single photograph that portrays a concept or a command. For example, pointing to a picture of a boy kicking a ball while saying “he kicks the ball” portrays the concept of a ball-kicking, and holding up a picture of a learner wearing his coat and saying “get your coat” communicates this directive. A scene cue is intended to visually capture a spoken message’s essence or intent; it is not an exact picture-for-word match to the spoken message, but an opportunity to process information as a whole (e.g., gestalt).

There are two types of scene cues: static and dynamic.

1. **Static scene cues.** A static scene cue uses an image or images that do not move, such as photos or drawings. There are two types of scene cues: “premade” and “buildable.” A premade scene cue is one in which a single image presents

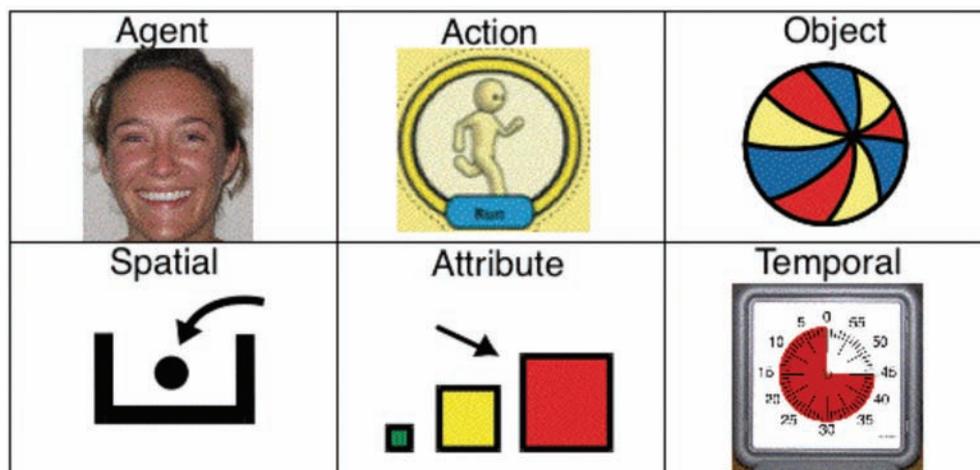


Figure 3-1. Examples of six types of element cues. (From *The Picture Communication Symbols* ©1981–2007 by Mayer-Johnson LLC. All Rights Reserved Worldwide. Used with permission.)

the entire scene, such as a picture of a boy riding a horse. A buildable scene cue combines more than one image to create a small scene. For example, separate pictures of a horse and a boy positioned in relation to each other against a background combine to represent the scene of a boy riding a horse.

- 2. Dynamic scene cues.** A dynamic scene cue uses an image that moves, such as an animated cartoon character that is running on a computer display, a video clip of a person running, or a toy figurine that can be manipulated by the learner or the instructor. For example, the instructor can sit down at a table with a flexible toy character, a car, a truck, and a block (Figure 3-2). To build the association between the visual scene and the action it represents, the instructor can say, “the boy pushes the truck,” “the boy climbs the ladder,” and “the boy jumps on the block” as the instructor demonstrates these maneuvers.

As the learner becomes accustomed to seeing visuals in a relevant context and in combination with speech, scene cues can be incorporated into daily life. At home, parents can take snapshots of commonly used items and then present these pictures at appropriate times. For example, a child’s mother might hold up a picture of the learner wearing his jacket as she says, “Time to get your jacket,” and then show a picture of the family in the car as she says “Let’s get into Mommy’s car.” Although it is not possible to have a visual for everything the child does, this technique is feasible and effective when parents focus on the actions related to the most common daily activities (such as getting ready for school, meals, and bedtime).

Communication Displays

Element cues and scene cues provide the content that fills in the target areas of the various communication displays



Figure 3-2. Examples of scene cues representing (1) Boy on block and (2) Boy push car.

instructors and parents use when working with learners with ASD. There are three types of displays: grid displays, whole scene displays, and mixed displays.

Grid Displays

The most common type of visual display features a grid format, that is, a blank background onto which the learner can place visual symbols, which typically ends up having a gridlike appearance. Most commonly, the grid is made of heavy paper or cardboard, and Velcro is typically used so the visual cues can be affixed to and removed from the grid. A computer screen also can function as a grid, often allowing for multiple pages; we will discuss computer displays in detail in Chapter 6. Common types of grid displays include:

Communication Book Displays

The grid format is used in communication books, which are notebooks containing graphic symbols. The first published descriptions of a communication book display format were designed to be used by learners with cerebral palsy who experienced limited or lack of speech (Goldberg & Fenton, n.d.; McDonald & Schultz, 1973; Vicker, 1974). The most well-known form of communication book display used by people with ASD is a PECS book (Frost & Bondy, 1994) that in use typically depicts preferred nouns representing desired items. The one difference between the original communication books and those used in the PECS program is a strip of removable Velcro on the cover of the book to which a learner places visual cues before handing the strip to the listener.

Conversation Displays

Grids can be used to display a series of symbols arranged by grammatical category that all relate to a single topic of particular interest to the learner, such as a favorite sports activity, his preferred games, information about a group of animals, or preparation of a favorite food. A “message window” is placed above the grid onto which visuals are affixed. For example, a grid display about preparing a pizza might include visuals of flour, tomato sauce, cheese, people involved in the preparation, utensils such as spoons and knives, and action verbs such as cutting, pouring, and baking (Figure 3-3). Conversation displays are used to aid comprehension as well as to assist in expression.

Conversation displays also can be used to remind the learner of a prior experience or direct him to carry out an activity. For example, an agent element cue of a picture of “Grandma” activates the experience of getting in the car, driving to Grandma’s house, and the anticipated experience at Grandma’s house. Combined action/object element cues such as the symbols for “drink + milk” can be used to direct the learner to pour himself a glass of milk.

Note that conversation displays are processed analytically—that is, the learner views each element one at a time and gradually gains meaning from the larger display based on his understanding of the semantic relationships between each element. Such analytic processing represents the beginning of the learner’s ability to generate a visual language system.

An important aspect of grid displays is that the initial grid is “decontextualized”—the display is out of the context in which it occurs. The learner begins



Figure 3-3. Example of a conversational display using multiple agents for preparing pizza. (From *The Picture Communication Symbols* ©1981–2007 by Mayer-Johnson LLC. All Rights Reserved Worldwide. Used with permission.)

with a blank slate, and the display gains meaning as items are added. In some cases, beginning with a decontextualized display may be advantageous, as it provides an open-ended forum for the learner to explore what he is most interested in. With conversational displays, the instruc-

tor should use whatever type of visual support the learner is most comfortable with, either from a communication book or from other sources, such as snapshots or pictures clipped from magazines. If possible, it is helpful to have some of the actual items available. To lay the basis for

30 VISUAL LANGUAGE IN AUTISM

broad use of language, the instructor should ask a variety of questions that will elicit many types of responses, such as: What is your favorite food? Who in your family likes ice cream? Which food did you eat yesterday?

If the learner requests a cookie, the instructor can expand the dialogue by asking if, for example, he wants to share the cookie. So ideally, instead of just saying, "I want a cookie," the individual learns sentences that use a variety of subjects, verbs, and objects, such as, "You're eating the cookie," and "I gave you a snack."

Use of conversation displays dates back to the 1940s, when Edith Fitzgerald proposed using them to help deaf children learn language structure and communicate (Fitzgerald, 1949. In Chapter 6, we provide greater detail on the creation of conversation displays.

Visual Schedules

Grids can be used to create a personalized daily schedule that lays out sequentially the important events of a learner's day (Figure 3-4). Visual symbols such as photographs, pictures, and text are arranged vertically or horizontally (MacDuff et al., 1993; Pierce & Schreibman, 1994). A similar display format can be used to outline a series of events or steps within a given activity, such as a class at school, a meeting with an instructor, or a birthday party. A more comprehensive description of the design and application of visual schedules is provided in Chapter 6.

Temporal Displays

Temporal displays help the learner understand time-related aspects of learning, such as how much time has elapsed, how

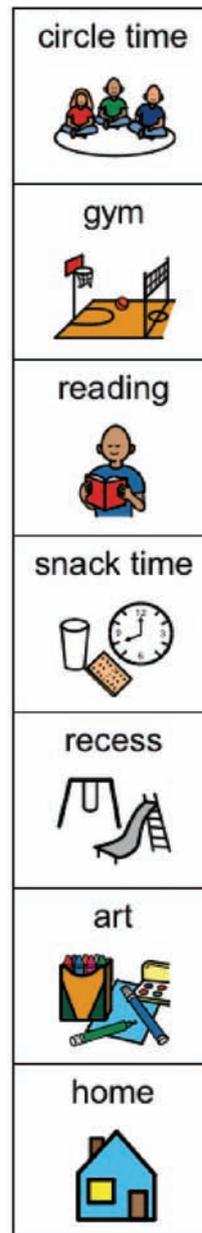


Figure 3-4. Example of a typical visual schedule. (From *The Picture Communication Symbols* ©1981–2007 by Mayer-Johnson LLC. All Rights Reserved Worldwide. Used with permission.)

many trials are needed to complete a task, and what the reinforcement for a specified activity is. Examples include:

- Visual timers, such as the Time Timer™, stopwatches, and hourglass timers
- First/then displays, such as a tabletop display that shows a visual representing a first activity (often the task) on the left side and a visual representing the follow-up activity (often the reward) on the right side. Chapter 6 provides greater detail on creating effective first/then displays.
- Countdown displays, such as boards with tokens removed each time the learner accomplishes a task, subtractive finger countdowns (such as the instructor holding up his hand and counting down each time the learner accomplishes a task), and reward

displays (such as the instructor presenting 10 pistachio nuts and allowing the learner to eat one for every correct answer).

Whole Scene Displays

Whole scene displays (also known as visual scene displays, or VSDs) provide a panoramic view of a large background. This enables the learner to portray events, people, actions, objects, and activities in the context within which they occur or exist. They can represent a generic context (e.g., a kitchen, a house with yard, or a school room with a teacher and students) or a personalized context (e.g., the view from a specific child's bedroom [Figure 3-5] or the view of a place he enjoys visiting, such as relative's house, the zoo, etc.).



Figure 3-5. Example of a whole scene display (playground).

32 VISUAL LANGUAGE IN AUTISM

Typically, a large photograph is used to depict the whole scene, and relevant objects can be added to the background by using Velcro. Computer screens also can be used; clicking on a “hot spot” in the background temporarily displays a larger picture.

As with scene cues, whole scenes can be “premade” or “buildable.” With a premade whole scene, the panoramic view already includes the key elements within it (i.e., a kitchen that includes the refrigerator, the stove, a microwave oven, a table, etc.). The learner can point out objects, place cut-outs on top of similar or identical objects, and add additional elements to the scene. With a buildable whole scene, the panoramic view is basically a background and the learner builds the scene by adding elements such as people or objects into it.

Whole scene displays can consist of a single scene that makes up the entire display, or multiple scenes on a page that reflect the place of occurrence, event, or sequence of the steps in a routine. To work with a whole scene, the learner and the instructor can sit at a table with the scene background displayed in front of them, along with relevant visual images. Visuals accompanying a whole scene depicting the view from the bedroom window might include a car, a squirrel, a dog, a ball, a snowman, a mailman, and other pedestrians. The instructor begins by asking a broad question, such as “What do you see outside your window,” and lets the learner’s answers guide the direction of the ensuing exchange.

In addition to skills the learner uses with conversation displays, whole scenes encourage expanded use of language, including greater use of verbs. For example, if the learner’s answer to the initial question is “I see a car,” the

instructor can encourage him to use “car” in sentences involving verbs such as “go,” “park,” and “drive.” As with grid displays, the instructor can ask the learner to act out some of the activities inspired by the scene. Questions such as, “Where do you kick the ball?” might lead to an exchange about balls, and the learner can be encouraged to say, “I kick the ball” as he does so in the room.

Whole scenes are an especially effective way to promote a learner’s ability to communicate, for three reasons:

- Whereas traditional picture books use images of isolated nouns to depict a person (*who*), place (*where*), or object (*what*) in isolation, whole scenes can convey all these themes simultaneously.
- Whole scenes are likely to provide greater meaning of schematic or experiential content to learners with ASD, including those who may not comprehend the semantic meaning behind images of isolated nouns.
- Whole scenes are particularly appealing to young children because they are realistic and personal, and their interactive nature tends to lead to enjoyable, engaging conversations. Light et al. (2006) reported that children 3 to 4 years of age had greater proficiency with information represented by a visual scene display than through a more traditional grid display with a more decontextualized symbol content.

Note that whole scenes are contextualized and processed as a unified whole (gestalt). The learner relies on the embedded context within the scene to glean meaning from it (Light et al., 2004).

Mixed Displays

Visual information can also be presented in ways that combine aspects of grid and scene displays. Such mixed displays can be used as a transitional bridge between

contextualized whole scenes and analytical conversation displays. For example, elements that describe the scene can be entered into a message window under the scene, as in Figure 3-6. Such a display combines analytical processing with



Figure 3-6. Mixed scene display depicting “boy brushes dog.” (From *The Picture Communication Symbols* ©1981–2007 by Mayer-Johnson LLC. All Rights Reserved Worldwide. Used with permission.)

gestalt processing, providing a structured method for gradually moving the learner from using whole scenes to using conversation displays.

The Three Modes

We now examine the three modes of the VIP and how the tools we have just covered can be used in these modes.

Visual Expressive Mode (VEM)

The first mode of the Visual Immersion Program involves using a full complement of visual resources for purposes of expressive communication. In the VEM, visual symbols such as line drawings, photographs, and entire visual scenes are used to express thoughts or ideas. The symbols may supplement a spoken message or stand alone. A person can use the VEM to make requests (e.g., for food, a pen, television, etc.), comment on feelings and experiences, give or get information, construct and answer questions, and so on. In its broadest sense, the VEM is a tool for a person to tell the story of his life.

Note that the VEM is not unique to individuals with autism. To understand this mode, it may be helpful to take a moment to consider that in all human social interaction, the output is expressive, and this output can be expressed with all types of symbols (spoken, gestural, visual, and graphic). Anytime we use a symbol for expression along with a visual element, we are using the VEM. Common examples of the VEM that we encounter in everyday life include handwritten letters, E-mail, close-captioning on

television, and the pictures and graphics used to support a newscaster's words.

The first reported use of visuals for expressive communication for individuals with ASD occurred in the early 1980s, with the recognition that children with ASD often have strong visual-spatial skills (Schuler & Baldwin, 1981). A host of visually based strategies followed, and using visuals for expression—especially for requesting desired goods and services—became a popular communication technique among parents, therapists and teachers who interact with individuals with autism. PECS, an innovation of the early 1990s, represents the most well-known example of using visuals to help learners with expression.

Using visuals to request items is highly effective because of the constant positive reinforcement inherent in the activity. The learner requests the items most important to him and then receives them, so once he learns to use visuals to make requests he is likely to continue doing so. This is an important first step. By making requests, the individual learns that symbols represent actual items, and he learns that he has the power to communicate.

It is important, however, to recognize that requesting is just a starting point in use of visuals for expression. Ideally, the learner should be able to do much more—to give and receive more detailed information, to ask questions, and to comment. In short, to engage in a conversation. Conversation displays, whole scenes, and scene cues can help learners express themselves more broadly.

Visual Instruction Mode (VIM)

The second mode of the VIP focuses on instruction—using visual elements (such

as photos, objects, or gestures) as a substitute or complement for spoken or written language to enhance commands or directions.

The emphasis of the VIM is on building language comprehension, which we view as the foundation for acquiring and applying language. In many cases, the learner may actually be capable of literally acting on a spoken command, but cannot process its intent because the linguistic structure is too advanced (a comprehension deficit) or the command is spoken too fast (an auditory processing deficit).

The simultaneous and sustained presentation of a visual that illustrates the ephemeral spoken command offers the learner a better opportunity for comprehension than a series of discrete visual symbols. (See *The Myth of a Visual Symbol Alignment* below.)

Aspects of the VIM are used in several existing approaches, including aided language (Goossens et al., 1992), aided language stimulation (Dexter, 1998), augmented input (Mirenda & Erickson, 2000; Wood et al., 1998), partner augmented input (Ronski & Sevcik, 1996),

The Myth of a Visual Symbol Alignment

When done improperly, constructing a statement or command from a series of visual images can create a confusing rather than clarifying experience. For example, mapping pictures to exact English word order can interfere with meaning because meaning comes across differently with discrete graphic symbols than it does through spoken words. What results is a string of disparate images whose meaning is largely impossible to decipher.

This phenomenon occurs because spoken language contains a rule-governed word order (syntax). A typically developing child learns this word order through repeated exposure to verbal communication during contextually based activities. Unfortunately, the rules of syntax do not directly transfer over to a display of visual symbols.

The problem is that constructing a sentence with a string of pictures focuses the child on the sentence's structure rather than meaning. For this technique to be effective, even for purely expressive purposes, the rule system regarding word order needs to be obvious and explicit to the learner.

Figure 3-A contains a set of pictures accompanied by English text for the sentence, "The boy pushes the girl who is sitting in the wagon." It is a perfectly coherent sentence in either spoken or written form because a rule-governed syntactic order with a specific semantic intent has been applied. A person who can read would have no trouble understanding the written words, and could match up the pictures to the relevant text.

However, consider what Figure 3-A conveys to someone who cannot read. If we eliminate the text, the message is nearly impossible to interpret, because the visual symbols have no obvious relationship to each other. The string of symbols is a series of discrete pictures that are not tied together by a set of linguistic conventions, which normally would improve the likelihood of comprehension.

Figure 3-B uses a scene cue to convey the same information contained in Figure 3-A. Although this scene cue is not governed by rules or part of an actual language, visual or otherwise, it clearly conveys the message, demonstrating that a simple picture express a complete idea.

Too often, we as practitioners have used approaches that place a series of symbols in exact English word order which we have come to believe is an intuitive but ultimately misguided attempt to address a complex problem. PECS books, computer-based communication programs, and many other communication approaches for individuals with ASD use discrete symbol elements that do not connect one grammatical category with another. Rather, these approaches use a set of graphic symbols that have meaning in isolation (e.g., a symbol for a watch looks like and can stand for a watch) or as isolated units, but do not necessarily have a logical association or relationship when those discrete symbols are arranged in sequence.

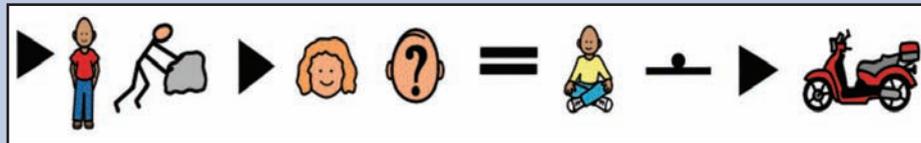


Figure 3-A. Example of one-to-one correspondence between visual symbols and spoken words in a sentence (to represent, “The boy pushes the girl who is sitting on the scooter.”) (From *The Picture Communication Symbols* ©1981–2007 by Mayer-Johnson LLC. All Rights Reserved Worldwide. Used with permission.)



Figure 3-B. Example of whole scene representing, “The boy pushes the girl who is sitting on the scooter.”)

and visually cued speech (Quill, 1995). These approaches tend to introduce symbols in an orderly manner, but the symbols are not typically presented in a way that makes language comprehension the primary focus.

Within the VIP scene cues are an essential component of the VIM. The advantage of scene cues is that they can represent a whole idea or complete thought through a single visual image. Eventually, the learner comes to understand what is portrayed in the scene cues: He understands the intention of the task by viewing the scene cue (generally accompanied by a spoken directive), and carries out the command.

From a developmental standpoint, scene cues serve two related purposes:

- They play a compensatory role, improving comprehension of a spoken message.
- They serve as a therapeutic tool to improve auditory comprehension, with the intention of gradually reducing dependence on the visual scene cue until a spoken message is understood on its own.

Because scene cues establish the link between a visual scene and the activity it represents, they are especially helpful for learners with ASD who are able to sort and match pictures and objects and can recognize nouns (characters) but have difficulty interpreting relational language (e.g., prepositions, action verbs). This disparate ability to comprehend different grammatical categories is common among learners with autism.

Scene cues also are useful for teaching practical tasks such as tying one's shoes or making a snack. The instructor displays the appropriate scene cue ("First we put

the right lace over the left lace," "Now we cut the sandwich in two pieces," etc.) as he or she demonstrates each step in the sequence of the task.

Recent research supports the use of scene cues: In numerous clinical trials where scenes cues were introduced, learners with ASD performed better when speech was combined with a scene than when speech was used alone (Shane & Douglas, 2002).

A key point about spoken language without accompanying visual support is that we often do not know which parts of the message the learner comprehends. He may understand all, part, or none of the spoken message, and all or part of the visual element. The important point is that combining the spoken and visual message is more effective than either component by itself. Exactly what percentage of information is understood through which avenue is less important than the goal of the learner understanding the overall intent of the message.

Note that the VEM and the VIM are often closely intertwined. A scene cue, for example, actually uses both modes. In making the statement, "Put the shoes in the closet" and displaying the visual, the instructor (the sender) is expressing information (use of the VEM). But from the perspective of the learner (the receiver), he is being instructed (use of the VIM).

The VIM and Sign Language

It should be pointed out that the VIM functions much like American Sign Language does for deaf individuals. Both communication methods do not rely on a literal word-for-word replication of spoken language. In other words, a sign language interpreter communicating the

Avoiding Unintended Interference to Comprehension Using Speech

In all three modes of VIP—and when working with learners with ASD in general—the instructor should avoid complex phrases and avoid rephrasing statements in a way that may confuse the learner.

This approach may run counter to the way adults often converse with typically developing individuals; if a toddler does not seem to understand what an adult says, the adult will often use different words to say the same thing. For example, a mother may hold up an apple and say, “Do you want an apple?” If the learner does not react, she may rephrase her question as, “Are you hungry—how about a snack?”

Although this technique is useful for typically developing children, it can make comprehension harder for individuals with ASD, who often take longer to process auditory information. As a result, rephrasing a question can have the unintended result of confusing the learner; as he attempts to process the words in the first phrase, a second sentence with different words arrives, and now all the words become jumbled, further muddying the sentence’s meaning.

We refer to this phenomenon as “unintended interference to comprehension.” To avoid it, we recommend repeating the identical phraseology, with sufficient pauses of several seconds between each attempt to give the learner who has a history of difficulty processing auditory information sufficient time to respond, before rephrasing the statement.

sentence “What time is your dentist appointment?” does not have six different signs for the words in this question, and an instructor using the VIM would not display six pictures to illustrate it.

Another similarity is that both communication methods may use multiple avenues of expression. A deaf person receives the information primarily from the speaker’s hand signs, but may also get some information from other cues, such as reading lips and viewing the speaker’s facial expressions. With the VIM, the learner has an opportunity to receive information from visual symbols in combination with spoken words.

Visual Organization Mode (VOM)

The final mode of the VIP is organization—using visual symbols to sequentially organize a task, activity, or schedule. Daily schedules and other time-oriented displays (timers, countdown displays, first-then displays) are common examples of the VOM, and this mode can also be used to depict the steps in a multipart task (such as making an omelet). A number of researchers have shown that visuals are effective for these purposes (Quill, 1995).

In all uses of the VOM, instructors should strive to give the learner as much

information as possible and to be as specific as possible. It is important to provide greater detail than is typically given when a single symbol is meant to represent an event (Shane, 2006). This additional information serves to offer clarity and reduce anxiety or confusion about the impending event or activity. For example, before beginning a 15-minute class on making a sandwich, the instructor can show the learner a display breaking the class down into three 4-minute segments on say, gathering the ingredients, making a sandwich, and eating the sandwich, with a 3-minute segment at the end for cleaning up. The individuals involved in the activity can be highlighted, as well as the activity location. Similarly, if a learner is being asked to complete a drawing task, the instructor might create a first-then display showing that when the learner draws (the “first” portion of the display), he can play a computer game (the “then” portion of the display) (Figure 3-7A). The VIP expands on this concept, whereby the opportunity to play for 10 minutes and create four drawings is depicted (Figure 3-7B).

Note that the information conveyed here is beneficial even when dealing with unpleasant things. First, the learner knows that the work portion (drawing four pictures) has a definite end, so he is less likely to become bored or frustrated during the activity. Second, he knows the fun portion (the computer game) has a definite end, so he is less likely to become upset when the endpoint arrives.

Behavioral Benefits: Order from Chaos

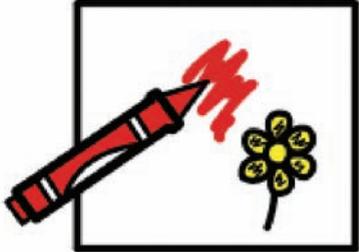
Used in this fashion, the VOM is an effective tool for helping individuals with ASD

reduce confusion about the day’s transitions and react to changes in their routine, minimizing anxiety and reducing both inward and outward aggression. It is not uncommon for a learner with autism to react negatively or even violently to changes in routine. A clear VOM serves as an alert about an impending transition and shows the learner what is about to happen. It can prevent a learner from being surprised when life’s normal transitions unfold.

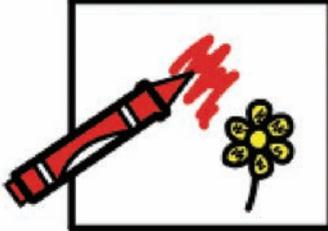
It is likely that the VOM leads to more socially appropriate behavior because behavior problems among individuals with ASD are often a secondary effect of a spoken language and information processing deficit. If a learner cannot comprehend basic spoken statements that mark transitions or denote abrupt and unexpected changes (e.g., “After you play the computer game for 10 minutes, you will have to go back to your regular class”), he has little way to know what is about to happen, and each event is a last-second surprise. The learner encounters similar time-oriented statements all day long (“Get your coat—we are going to Grammy’s house,” “Time to turn off the TV—it’s time for dinner”), so it is vital that he learn to understand them.

Not having the ability to understand such instructions is a little like being awoken from the comfort of a sound sleep and moved to a new location. For the learner with a spoken language difficulty who is uncertain about the events of the day, this hasty awakening occurs constantly, making life a chaotic experience. Using the VOM to give a warning often has an immediate and positive effect on behavior, attitude, and learning.

The VOM also is an important tool for establishing routines and habits,

First	Then
	

A.

First		Then	
			
			

B.

Figure 3-7. **A.** Example of a typical first/then display (representing “first draw, then computer”). **B.** Example of an expanded first/then display (representing “first draw 4 pictures, then computer for 10 minutes”). (From *The Picture Communication Symbols* ©1981–2007 by Mayer-Johnson LLC. All Rights Reserved Worldwide. Used with permission.)

which people with ASD often find deeply comforting. The VOM can be used to create a display showing after-dinner activities, such as the individual getting to watch a video after he straightens up his room.

More broadly, the VOM plays a key role in laying the foundation for a more complex understanding of time. Comprehending that certain events have a causal relationship with other events helps the learner develop the ability to predict and infer what is going to happen based on prior experiences. This is critical for the learner to come to understand that rewards are not always immediate. For example, it is an important behavioral step when a learner can understand that if he behaves well for a month, he will get to go to a baseball game, or that if he completes certain household chores, he will get a monthly allowance.

Does Visual Language Enhance or Blunt Speech?

In practice, the VEM, the VOM, and the VIM are used together. Over time, the learner becomes comfortable using visual scene displays, visual schedules, scene cues, and other tools from the visual arsenal to communicate. This raises an important issue, which is sometimes voiced by parents and instructors: Does using visuals as the foundation for communication interfere with spoken language acquisition?

Based on our clinical experience, the answer is no. In fact, we believe just the opposite: We consider visuals an excellent form of speech therapy. We have worked with many individuals with ASD

who began speaking or made significant improvements in their speaking abilities after they learned to use visual language.

We do not know precisely why this is the case. Perhaps seeing visual images allows learners with ASD to better organize information and rehearse what they want to say. At a neurological level, perhaps visual language activates parts of the brain that were not previously used for communication (the phenomenon of “intersystemic reorganization”) (Luria, 1980).

Regardless of the mechanism, the important point is that visual speech enhances rather than blunts the learner’s ability to speak. We consider it analogous to crutches—people who break a leg only use crutches as long as they need to; they are anxious to start walking again as soon as possible. Similarly, individuals with ASD want to speak, and they will not rely on visuals to communicate if they have the ability to use spoken language.

The level of speech development varies by learner. Regrettably, some will never develop sufficient speech skills to become what we would consider competent speakers. Some will use a combination of visuals and spoken words or phrases. Some individuals make so much progress that they no longer need visuals to communicate. Regardless of the scope of their progress, we believe that visual language enables learners to reach their full potential for speech.

By now, readers should have a general idea of how the VEM, the VOM, and the VIM can be used to help learners with ASD. We go into greater detail on the practical aspects of applying these modes in Chapter 6. First, however, we describe how the VIP can be used as an assessment tool in Chapters 4 and 5.

Three Therapy Models

The complexity of the communication problems associated with the ASD population demand that instructors conceptualize therapeutic approaches to communication disorders in three ways. The *fix-it* model aims to correct the underlying source of the learner's disorder. The *compensation* model aims to give the learner alternative ways to handle the existing disorder. The *bypass* model circumvents the disorder with an alternative approach.

The Visual Immersion Program may follow all three models, depending on the learner's developmental capabilities. It is important to continually assess the learner's response to intervention and modify the therapeutic approaches as appropriate.

Fix-it model. For some individuals with ASD, the VIP serves as a remedial tool that strongly supports the development of receptive and expressive speech. Until the advent of augmentative communication approaches for learners with ASD this was the exclusive model of intervention. Understandably, this is the outcome that all parents and instructors hope for.

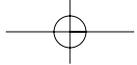
Compensation model. Most commonly, visuals serve to complement, support, and enhance the learner's ability to understand and speak in effect compensating for their communication deficits. In some cases, the continued supplementation offered by this approach may lead to comprehension and production that does not require accompanying visual supports. In fact, for some people, there is a steady and predictable improvement in communication.

Bypass model. For individuals who are unable to develop the ability to speak, the VIP provides an alternative way to communicate, bypassing traditional speech in the same way sign language bypasses speech for a deaf person. The bypass model is an example of the purest form of augmentative communication.

Regardless of which model the VIP most closely follows for a particular learner, instructors and parents should always bear in mind that the program's goal is to enable the learner to communicate and function as effectively as possible within his environment. It is understandable that many parents will hope for a complete fix, but they should be prepared if the therapy ultimately serves to compensate for the learner's speech difficulty or bypass it entirely. In every case the goal of intervention is to provide the individual with ASD the most appropriate and richest communication system available to him.

References

- Bondy, A. S., & Frost, L. (1994). The Picture Exchange Communication System. *Focus on Autistic Behavior, 9*, 1-9.
- Bondy, A. S., & Frost, L. A. (1998). The Picture Exchange Communication System. *Seminars in Speech and Language, 19*, 373-398.
- Charlop-Christy, M., Carpenter, M., Le, L., LeBlanc, L., & Kellet, K. (2002). Using the Picture Exchange Communication System (PECS) with children with autism: Assessment of PECS acquisition, speech, social-communicative behavior, and problem behavior. *Journal of Applied Behavioral Analysis, 35*(3), 213-231.
- Dexter, M. (1998). *The effect of storybook aided language stimulation on the communicative output of children with PDD-NOS*. Unpublished doctoral dissertation, Johns Hopkins University, Baltimore.
- Fitzgerald, E. (1949). *Straight language for the deaf*. Washington, DC: The Volta Review.
- Frost, L., & Bondy, A. S. (1994). *PECS: The Picture Exchange Communication System training manual*. Cherry Hill, NJ: Pyramid Educational Consultants.
- Goldberg, H. R., & Fenton, J. (Eds.). (n.d.). *Aphonic communication for those with cerebral palsy*. New York: United Cerebral Palsy Association.
- Goossens, C., Crain, S. S., & Elder, P. S. (1992). *Engineering the preschool environment for interactive, symbolic communication*. Birmingham, AL: Southeast Augmentative Communication Conference Publications.
- Light, J., Drager, K., McCarthy, J., Mellott, S., Parrish, C., Parsons, A., et al. (2004). Performance of typically developing four and five year old children with AAC systems using different language organization techniques. *Augmentative and Alternative Communication, 20*, 63-88.
- Luria, A. R. (1980). *Higher cortical functions in man*. New York: Basic Books.
- MacDuff, G. S., Krantz, P. J., & McClanahan, L. E. (1993). Teaching children with autism to use photographic activity schedules: Maintenance and generalization of complex behavior chains. *Journal of Applied Behavior Analysis, 26*, 89-97.
- McDonald, E., & Schultz, A. (1973). Communication boards for cerebral palsied children. *Journal of Speech and Hearing Disorders, 38*, 73-88.
- Mirenda, P., & Erikson, K. A., (2000). Augmentative communication and literacy. In A. M. Wetherby & B. M. Prizant (Eds.), *Autistic spectrum disorders: A transactional developmental perspective* (pp. 333-367). Baltimore: Paul H. Brookes.
- Pierce, K., & Schriebman, L. (1994) Teaching daily living skills to children with autism in unsupervised settings through pictorial management. *Journal of Applied Behavioral Analysis, 27*, 471-482.
- Quill, K. A. (1995). Visually cued instruction for children with autism and pervasive developmental disorders. *Focus on Autistic Behavior, 10*(3), 10-20.
- Romski, M., & Sevcik, R. (1996). *Breaking the speech barrier: Language development through augmented means*. Baltimore: Paul H. Brookes.
- Schuler, A., & Baldwin, M. (1981). Nonspeech communication and childhood autism. *Language, Speech and Hearing Services in the Schools, 12*, 246-257.
- Shane, H. C. (2006). Using visual scene displays to improve communication and communication instruction in persons with Autism Spectrum Disorders. *Perspectives in Augmentative and Alternative Communication, 15*(1), 8-13.
- Shane, H. C., Douglas M. L. (2002, November). *Investigation into the use of intelligent agents in children evidencing autism*. Presentation at the annual meeting of the American Speech-Language-Hearing Association; Chicago, IL.



44 VISUAL LANGUAGE IN AUTISM

Vicker, B. (1974). *Nonoral communication system project: 1964/1973*. Iowa City: Campus Stores, Publishers, University of Iowa.

Wood, L., Lasker, J., Siegel-Causey, E., Beukel-

man, D., & Ball, L. (1998). An input framework for augmentative communication. *Augmentative and Alternative Communication, 14*, 261-267.

